UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

JOHN DOE,

Plaintiff,

v.

COLUMBIA UNIVERSITY,

Defendant.

No. 1:20-cv-06770-GHW

STIPULATION REGARDING
ELECTRONIC DISCOVERY AND ORDER

The Parties in the above-captioned litigation (the "Action") hereby stipulate to the following procedures, protocols and obligations for the location, collection and production of hard-copy documents and electronically stored information ("ESI") by any Party in or pursuant to the Action.

I.   DEFINITIONS

A. "Bates Number" has the meaning given in Section III.I of this Stipulation and Order ("Stipulation").

B. "Custodian" means the Person who is responsible for, has administrative control over, or has access to potentially responsive ESI.

C. "Discovery Material" includes all information exchanged between the Parties, whether gathered through informal requests or communications between the Parties or their counsel or gathered through formal discovery conducted pursuant to Rules 30 through 36. Discovery Material includes information within documents, depositions, deposition exhibits, and other written, recorded, computerized, electronic, or graphic matter, copies, and excerpts or summaries of documents disclosed as required under Rule 26(a).

D. "Document" is defined as set forth in Local Civil Rule 26.3(c)(2)—that is: (1) "synonymous in meaning and equal in scope to the usage of the term 'documents or electronically stored information' in Fed. R. Civ. P. 34(a)(1)(A)"; and (2) "a draft or non-identical copy is a separate document within the meaning of this term." For the avoidance of doubt, the term "Document" as used herein includes Hard-Copy Documents and ESI.

E. "Electronic Communication" means a message transmitted electronically or digitally and includes emails, text messages, multimedia messages, communications from any social media platform (whether public or private posts or private messages) or messages from any chat application or chat feature (including Microsoft Teams, Slack, Skype, Zoom, Signal, Hangouts, Messenger, and WhatsApp), as well as any file (document, photo, video, audio, or otherwise) incorporated as an attachment or embedded file or sent or received independently (that is, not accompanied by an electronic message containing text).

F. "Electronically Stored Information" or "ESI" is equal in scope to "electronically stored information" in Fed. R. Civ. P. 34(a)(1)(A) and includes information or data (including files, databases, and programs) stored or recorded in or on electronic, digital, magnetic, or similar Media. For the avoidance of doubt, ESI includes Electronic Communications.

G. "Extracted Text" means text extracted from a Native Format file using standard e-discovery processing tools.

H. "Family Group" has the meaning given in Section III.G of this Stipulation.

I. "Family Relationship" means the association between an attachment and its parent document or between embedded documents or linked internal or non-public documents and their parents.

J. "Hard-Copy Document" means a document existing in paper or other tangible form at

the time of collection.

K. "JPEG" refers to the Joint Photographic Experts Group file format.

L. "Load File" means the electronic file that is used to import all required production information into a document database, including, as applicable: document images; Extracted Text or Optional Character Recognition ("OCR") Text; Native Format files where required by this Stipulation; Metadata; information related to the embedded content of a document; and information indicating document breaks and document relationships, such as the relationship between an Electronic Communication and its attachments.

M. "Local Civil Rules" means the Local Civil Rules of the United States District Courts for the Southern and Eastern Districts of New York, effective October 15, 2021.

N. "Media" means an object or device, real or virtual, on which information or data is stored, including but not limited to digital or analog machine-readable devices, computers, Internet sites, portable storage media, servers, networks, cloud-based storage, discs, and tapes.

O. "Metadata" means (i) structured, i.e., fielded, information embedded in a Native Format file which describes, *inter alia*, the characteristics, origins, usage, and/or validity of the electronic file; and (ii) information generated automatically by the operation of a file system or application when a native file is created, modified, transmitted, deleted, or otherwise manipulated by a user of such system.

P. "Native Format" means the file format of ESI in the application in which such ESI was originally created.

Q. "Optical Character Recognition" or "OCR" means the process of recognizing, and creating a file containing searchable text (the native text extracted or generated from a Document) within an image of a document. The latter text is also referred to as "OCR Text."

R. "Party" means any party to the Action, as well as its or their directors, principals, employees, and agents.

S. "Person" is defined as set forth in Local Civil Rule 26.3(c)(6)—that is, "any natural person or any legal entity, including, without limitation, any business or governmental entity or association."

T. "Privileged Information" means Documents protected from production or disclosure (in whole or in part) by the attorney-client privilege, attorney work product protection, common-interest privilege, or any other legal privilege or immunity.

U. "Privilege Log" has the meaning given in Section V.D of this Stipulation.

V. "Producing Party" means a Party that produces Discovery Material in this Action.

W. "Production Media" has the meaning given in Section III.L of this Stipulation.

X. "Protective Order" means any protective order governing the disclosure of confidential Discovery Material entered by the Court in this Action (and any subsequent amending orders thereto).

Y. "Requesting Party" means any Party requesting Discovery Material through written discovery.

Z. "Receiving Party" means a Party that receives Discovery Material from a Producing Party in this Action.

AA. "TIFF" or Tagged Image File Format refers to the CCITT Group IV graphic file format for storing bit-mapped images of ESI or scanned Hard-Copy Documents.

II.   GENERAL PROVISIONS

A. Construction of Certain Terms in this Stipulation. The rules of construction detailed in Local Civil Rule 26.3(d) shall apply to the construction of this Stipulation.

B. Confidentiality. Discovery Material shall be subject to the terms of the Protective

4

Order, as well as Federal Rule of Civil Procedure 26(b)(5)(B).

C. <u>Preservation of Objections</u>. By entering into the Stipulation, no Party waives any right it otherwise might have to object to disclosing or producing any information or item on any ground, including confidentiality, attorney-client privilege, attorney work product doctrine, or any other applicable and legally cognizable protection, privilege, or immunity. Similarly, no Party waives any right to object on any ground to the discoverability, authenticity, admissibility, or use in evidence of any of the material covered by this Stipulation.

D. <u>No Modification to Retention Obligations</u>. Nothing in this Stipulation alters a Party's legal obligation to retain Documents and data.

E. <u>Exception Files</u>. The Parties will use reasonable efforts and standard industry practices to address Documents that present imaging or form production problems (including encrypted and/or protected files identified during the processing of ESI, and rendering to TIFF images) ("<u>Exception Files</u>"). The Producing Party shall produce a log containing identifying Metadata and file contents that can be extracted from the Exception File ("<u>Exception Log</u>"). If a Party requests the production of any Exception File listed on the Exception Log, the Parties will meet and confer regarding whether a reasonable and cost-effective means exists for obtaining and producing the requested Exception File.

F. <u>Costs</u>. Without prejudice to any application for costs or award of a reasonable attorney's fee, or to any subsequent agreement between the Parties regarding the allocation of costs, each Party agrees to bear its own costs of locating, maintaining and producing Documents in this Action. The Parties expressly reserve their rights to seek their costs in connection with any dispute the resolution of which is unaddressed by or departs from the terms of this Stipulation, and to seek recovery of a reasonable attorney's fee incurred in connection with discovery to the extent

permitted under 42 U.S.C. § 1988(b).

G. <u>Disputes</u>. The Parties acknowledge their obligations under Local Civil Rule 37.2 and Section 2.C.ii of the Individual Rules of Practice in Civil Cases of Gregory H. Woods, including to attempt to confer in good faith in an effort to resolve any discovery dispute prior to seeking judicial resolution of such a dispute. The Parties agree that the obligations under Local Civil Rule 37.2 and Section 2.C.ii extend to any disputes that arise under or in relation to this Stipulation. As such, in the event the Parties cannot reach agreement on a disputed matter, the Parties may submit the matter to the Court in accordance with the procedures and requirements outlined in Local Civil Rule 37.2 and Section 2.C.ii of the Individual Practices in Civil Cases of Gregory H. Woods.

III.   PRODUCTION FORMAT

A. <u>Format Guidelines</u>. The Parties shall, to the extent reasonably and technically possible, produce Documents according to the specifications provided in this Stipulation and in Exhibit A.

B. <u>De-NISTing</u>. ESI shall be filtered for file type to remove system files, program files, and other non-user created data files that are on the industry standard list of such files maintained by the National Institute of Standards & Technology ("NIST"). The Parties recognize that to reduce the document review population, additional file types may need to be excluded.

C. <u>Deduplication</u>. A Party is required to produce only a single copy of a Responsive Family Group. Each Party may remove exact duplicate Family Groups (i.e., identical copies of the same Document(s)), to reduce the unnecessary cost of reviewing and producing exact duplicate Documents. If a Party chooses to remove exact duplicate Documents, the Party shall make reasonable efforts to remove exact duplicate ESI according to the MD5 or SHA-1 hashing method and additional information regarding the Family Relationship and shall identify in the All Custodians field of each Document the name(s) of the Custodians whose duplicate Documents the Producing Party removed during processing. Moreover: (1) if a Document is a duplicate of another

Document but the Documents have different family members, they will not be deduplicated,[1] although each family member shall be hashed separately for purposes of populating the "MD5 or SHA-1 Hash" field in Exhibit A; (2) responsive non-privileged attachments to Documents shall not be disassociated from the parent Document even if they are exact duplicates of another Document in the production; and (3) Hard-Copy Documents shall not be eliminated as duplicates of responsive ESI. ESI that is not an exact duplicate according to the MD5 or SHA-1 hashing method may not be removed. However, a Producing Party may use email threading—that is, to enable the Producing Party to produce the most inclusive subset of emails in a responsive population—to defensibly exclude, from both review and production, duplicative content within the Producing Party's Documents. Both deduplication and email threading shall be performed with careful consideration of full Family Groups, and as such, neither process will suppress unique content found in an attachment. The Producing Party shall produce separate inclusive emails where a thread splits into different conversation "branches."

D. Encrypted Documents. The Parties will make reasonable efforts to ensure that any responsive Documents identified that are encrypted or password-protected are successfully processed for review and production under the requirements of this Stipulation, and if produced in Native Format, the decrypted Document is produced. To the extent responsive encrypted or password-protected Documents are successfully processed according to the requirements of this Stipulation, the Parties have no duty to identify the prior encrypted status of such Documents. To the extent such responsive Documents are not successfully processed despite use of reasonable

---

[1] Unless otherwise agreed by the Parties, a Family Group can only be removed through deduplication if each and every member of the deduplication candidate Family Group has identical MD5 or SHA-1 hash codes to the corresponding members of another Family Group. For example, if there is a two-member Family Group with a parent and an attachment, the MD5 hash codes of the parent must be identical and the MD5 hash codes of the attachments must also be identical.

efforts, including reasonable technical efforts to obtain passwords, a placeholder TIFF image may be produced in place of each such Document indicating that security protection could not be removed. Upon request from the Receiving Party, the Parties shall meet and confer in good faith regarding reasonable efforts or mechanisms that could be undertaken to remove such security protection or the production of available Metadata.

      E. <u>Documents to Be Produced in Native Format</u>. As provided in Exhibit A:

      1. Microsoft Excel and other spreadsheet files, including comma or tab delimited text files, photographs, video files, audio files, animation files, and other multimedia files shall be produced in Native Format, if available.

      2. Photographs, PowerPoint, and other presentation files may be produced either in Native Format or as color images in TIFF format.

      3. The Parties agree to meet and confer regarding the ability to produce other file types in Native Format.

      4. If a Document to be produced in Native Format contains Privileged Information or information subject to any other applicable protection, the Document will be produced either in Native Format with redactions or by producing the Document in TIFF format with redactions, in accordance with Section V of this Stipulation.

      5. The production Load Files shall contain a link to the produced Native Format files as specified in the "NativePath" Metadata field described in Exhibit A. Each electronic file produced in Native Format shall be assigned a unique Bates Number, as set forth in Section III.I, and the database record for that file shall include a single page Bates-stamped TIFF image slip-sheet bearing the phrase "PRODUCED IN NATIVE FORMAT" (or similar language). To protect the confidentiality of files produced in Native Format, any confidentiality designations pursuant

to the Protective Order shall appear on the associated TIFF slip-sheet. Files produced in Native Format shall be given file names that include the Bates Number and the file extension.

F. <u>Production of Additional Documents in Native Format</u>. Each Party reserves the right to request production in Native Format for any ESI that it believes is not adequately represented in the TIFF format specified in Exhibit A and, upon a showing of reasonable need, such requests will not be unreasonably denied.

G. <u>Family Relationships and Family Groups</u>.

1. A Document and all other Documents in its attachment range, emails with attachments, and email, Electronic Communications, or other Documents together with any Documents referenced by Document stubs or via links to internal document sources within those emails, Electronic Communications or other Documents all constitute "<u>Family Groups</u>".

2. If any member of a Family Group is produced, all members of that Family Group must be also be produced, except (1) any Documents withheld or redacted on the basis of privilege or other applicable legally cognizable protection; and (2) any nonresponsive attachments. If the Producing Party withholds or redacts a Document under (1), the Producing Party shall comply with Section V of this Stipulation. If nonresponsive attachments are withheld under (2), the Producing Party must produce a placeholder TIFF image in place of each such Document indicating that the Document is nonresponsive.

3. All Documents within a Family Group shall be produced together. Family Group relationships shall be preserved. Child files shall follow parent files in sequential Bates Number order, including in the case of Embedded Files produced separately from their parent.

H. <u>Scanning Hard-Copy Documents and Unitization</u>. In scanning Hard-Copy Documents, each page should be output to a single page TIFF file. Hard-Copy Documents should be scanned

as they are kept. Distinct, logical document breaks should be defined as such in a standard Load

File as described in Exhibit A. The Parties shall make reasonable efforts to unitize the Documents

correctly (i.e., distinct documents should not be merged into a single record, and a single document

should not be split into multiple records). The Parties shall also make reasonable, best efforts to

maintain the unitization of distinct documents within an organized group structure, such as folders,

clipped or stapled bundles, and binders. In the case of an organized compilation of separate

Documents (e.g., a binder containing several separate Documents behind numbered tabs), the

Document behind each tab should be scanned separately, but any Family Relationship among the

scanned Hard-Copy Documents in the compilation should be reflected in the data Load File at the

appropriate standard fields. Where a Document or organized group structure has a cover, spine,

tab, divider, or any other identifying label, the information from that Document or organized group

structure shall be scanned and produced as the first page of the Document or Family Group. Pages

containing post-it notes or other detachable notes that obscure the underlying Document should be

scanned once with the detachable note intact, and then again without it, and made part of the same

Document.

I.   Bates Numbers. Each page of a Document produced in TIFF format and each electronic

file produced in Native Format shall have a unique fixed-length numeric identifier ("Bates

Number") containing at least eight (8) digits. The Bates Number for each page of a Document shall

be created so as to identify the Producing Party, including a unique alpha prefix for each individual

Named Plaintiff, and the Bates Number (e.g., "RICH_00000001" or "NYC_00000001"). The

Producing Party may choose its own alpha prefix but that prefix shall not change across

productions. Child files shall follow parent files in sequential Bates Number order. Bates Numbers

shall be electronically "burned" onto the image following the formatting requirements in

Exhibit A. Unless it would obscure, conceal, or interfere with any information originally appearing in the underlying image, the Bates Number shall be burned on the lower right-hand corner of each image.

J.   Confidentiality Designations for TIFF Images. Unless it would obscure, conceal, or interfere with any information originally appearing in the underlying image, any confidentiality designation pursuant to the Protective Order shall appear on the lower left-hand side of each image.

K.   Metadata Fields and Processing. Each of the Metadata and coding fields set forth in Exhibit A that can reasonably and technically be extracted from a Document shall be produced for that Document. The Parties are not obligated to populate manually any of the fields in Exhibit A if such fields cannot be extracted from a Document using reasonable means, with the exception of the following, if available: (a) BegBates, (b) EndBates, (c) BegAttach, (d) EndAttach, (e) Custodian, and (f) MD5 Hash or SHA-1 Hash. Metadata shall be provided in a Concordance-format delimited file with a .DAT file extension and standard delimiters for column break and text qualifier. The first line shall be the header with field names, and each subsequent line shall contain the fielded data for each Document. Nothing herein shall require a Producing Party to create or produce Metadata that does not exist or is not reasonably or technically accessible.

L.   Production Media. The Producing Party shall produce Document images, Native Format files, Load Files, and Metadata on hard drives, CDs, DVDs, USB thumb drives, secure FTP, or other mutually agreeable Media ("Production Media"). Each piece of Production Media shall include a unique identifying label corresponding to the identity of the Producing Party, the date of the production of Documents on the Production Media, and the Bates Number ranges of the Documents in that production (e.g., "Defs.' Production January 1, 2022, NYC_00000123-NYC_00000456"). To the extent that the Production Media includes any confidential information

protected under the Protective Order, the label on such Production Media shall indicate that the

Production Media includes information so designated as required under the terms of such order.

Production Media shall also include text referencing the case name and number. Further, any

replacement Production Media shall clearly identify that it is a replacement and cross-reference

the Bates Number range that is being replaced. All Production Media that is capable of write

protection should be write-protected before production. All Production Media shall be encrypted,

with the Producing Party to provide a decryption key at the time of production. Productions should

be accompanied by a cover letter identifying the Bates Numbers for Documents produced. All

electronic transmissions of produced materials must be effected via a secure file transfer system.

IV.   PROCESSING SPECIFICATIONS

A. The Producing Party will generate and preserve the MD5 or SHA-1 hash values of all

ESI based on the native file.

B. The Producing Party shall use the following specifications when converting ESI from

its Native Format into TIFF image files prior to its production: Where Documents are not otherwise

produced natively, all TIFF images shall display "track changes" (and similar in-line editing),

comments, PowerPoint speaker notes, and other hidden text, regardless of the display setting for

this information as last saved by the Custodian, to the extent reasonably and technically possible.

C. Each Producing Party shall make reasonable efforts to process all ESI with the

following time zone: Eastern / GMT-5.

V.   REDACTED AND WITHHELD DOCUMENTS

A. Redaction and Withholding of Certain Documents. If a claim of privilege or other

legally cognizable protection applies to only a portion of a Document, the Document should be

produced (in Native Format or TIFF image) with the portion claimed to be privileged or protected

obscured and the redacted area stamped with a label containing the reason for the redaction, for

example "Redacted—Privileged" or similar. If a Producing Party wishes to redact Native Format photo, video, or audio files for reasons other than to protect the privacy of individuals not party to this Action, the Parties shall meet and confer as to the bases for such redaction, the format of redaction to be applied, and the format of production for the unprotected portion of the content. Any redacted material must be clearly labeled as having been redacted. To the extent responsive Documents contain privileged or protected information and cannot be redacted or produced in TIFF format, such Documents may be wholly withheld.

B. <u>Redaction of Certain Non-Responsive Information</u>. A Producing Party may redact irrelevant and/or non-responsive information.

C. <u>Efficient Communication of Privilege Claims</u>. The Parties agree to meet and confer in good faith as to the use of categorical Privilege Logs for specific categories of Documents and any other circumstances in which departure from the use of traditional document-by-document privilege logs would further the just, speedy, and inexpensive determination of the Action.

D. <u>Privilege Logs</u>. The Producing Party shall provide the Receiving Party with a log of the documents withheld or redacted for privilege or any other basis, consistent with Federal Rule of Civil Procedure 26(b)(5), Local Civil Rule 26.2, and any other applicable law ("<u>Privilege Log</u>"). Privileged communications with litigation counsel that occurred after July 1, 2020, need not be logged.

E. <u>Metadata for Redacted Documents</u>. The Parties agree that for redacted documents, all Metadata fields can be withheld except: BegBates; EndBates; AttachRange; BegAttach; EndAttach; and PageCount.

F. <u>Content of Privilege Log for Withheld Documents</u>. If a document-by-document privilege log is used, the Party claiming the privilege shall state, at minimum, the following

information separately for each Document withheld:

1. Custodian, From, To, CC, BCC, Bates Number, Date;

2. where not apparent, the relationship of the author, addressees, and recipients to each other, designating counsel names with an asterisk;

3. the nature of the privilege (including work product) or other legally cognizable protection which is being claimed;

4. the type of Document; and

5. the general subject matter of the Document.

G. <u>Content of Privilege Log for Redacted Documents</u>. The Producing Party shall indicate the basis for the redaction on a separate log, providing the Bates Number of the redacted Document, the applicable claim(s) of privilege or other legally cognizable protection, and any of the other information called for above in (1) through (5) that is not evident from the face of the redacted Document. If there is more than one reason for a redaction, each reason shall be addressed in the log with sufficient information to allow the Receiving Party to understand the basis for such a redaction.

H. <u>Additional Formatting and Substantive Requirements</u>. The Parties shall meet and confer in good faith regarding any additional substantive or formatting requirements for Privilege Logs.

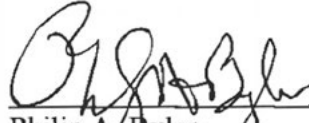## VI.   **VARIATION OF STIPULATION'S REQUIREMENTS**

Nothing in this Stipulation shall preclude any Party from seeking to amend any procedure or requirement set forth herein. Any procedure or requirement of this Stipulation may be varied by agreement of the Parties, confirmed in writing, where such variance is deemed appropriate to facilitate the timely and economical exchange of ESI or other covered Discovery Material.

IT IS HEREBY AGREED by the undersigned as of May 15, 2023.

Roberta A. Kaplan
Gabrielle E. Tenzer
KAPLAN HECKER & FINK LLP
350 Fifth Avenue, 63rd Floor
New York, NY 10118
Telephone: (212) 763-0883
Facsimile: (212) 937-3734
rkaplan@kaplanhecker.com
gtenzer@kaplanhecker.com

Attorneys for Defendant
The Trustees of Columbia
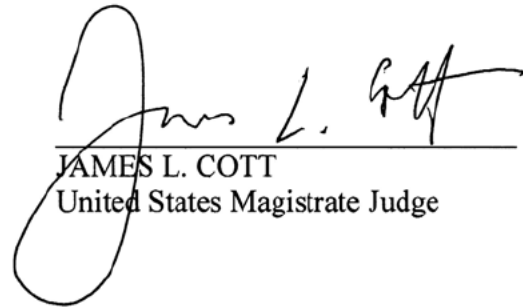*University in the City of New York*

Philip A. Byler
Andrew T. Miltenberg
NESENOFF & MILTENBERG, LLP
363 Seventh Avenue, Fifth Floor
New York, NY 10001
Telephone: (212) 736-4500
pbyler@nmllplaw.com

Attorneys for Plaintiff John Doe

SO ORDERED

Dated: ___05/19/2023___
New York, New York

JAMES L. COTT
United States Magistrate Judge

15

EXHIBIT A

DOCUMENT PRODUCTION FORMATS

I.   **OVERVIEW**

A. All Documents shall be produced: (1) as (i) Bates-stamped tagged image file format ("TIFF") images, or (ii) where expressly required or permitted by this Exhibit, Native Format files compliant with the Native Format requirements detailed Section IV of this Exhibit; and (2) along with (i) an image load/cross reference file, (ii) a data Load File with fielded Metadata, and (iii) document-level Extracted Text for ESI or OCR Text for scanned Hard-Copy Documents.

B. All produced Documents shall comply with the applicable requirements set out in this Exhibit and in the Stipulation.

II.   **IMAGE REQUIREMENTS**

A. Except as otherwise provided in this Stipulation, all Documents shall be produced as TIFF images in 300 dpi Group IV single-page monochrome format.

B. Scans of Hard-Copy Documents or redacted ESI that contain color that is reasonably necessary to decipher the meaning, context, or content of the Document (e.g., graphs or embedded images; redlines) shall be produced as single-page, 600 DPI JPEG images with JPEG compression and a high quality setting as to not degrade the original image.

C. All images shall be sequentially Bates-stamped, in accordance with the requirements set out in Section III of this Exhibit.

D. Images shall include the following content where present:

1.   For word processing files (e.g., Microsoft Word) – Comments and "track changes" (and similar in-line editing).

2.   For presentation files (e.g., Microsoft PowerPoint) – Speaker notes and comments.

E.   Reasonable care shall be taken not to degrade the legibility of Documents during the imaging process. If legibility is found to have been degraded, the Receiving Party may make reasonable requests for re-production of affected Documents.

F.   If particular Documents warrant departure from these image requirements (e.g., production of color images instead of monochrome), the Parties shall cooperate to arrange for the mutually acceptable reasonable production of such Documents.

III.   **BATES STAMPING REQUIREMENTS**

A.   Each page of a Document produced in image format shall have its Bates Number electronically "burned" onto the image, using a consistent length with leading zeros in the number, and consistent font and type size.   The format of the Bates Numbers shall not change across productions.

B.   Unless it would obscure, conceal, or interfere with any information originally appearing on the Document, the Bates Number shall be burned on the lower right-hand corner of the Document.

IV.   **NATIVE FORMAT REQUIREMENTS**

A.   Spreadsheet files

1.   Spreadsheet files (e.g., Microsoft Excel) shall be provided in Native Format.

2.   In lieu of a TIFF image version of each spreadsheet file, a Bates-stamped single-page TIFF placeholder file shall be produced along with the Native Format version of each file. The TIFF Placeholder should read "Document Produced Natively."

3.   Spreadsheet files shall include the following content where present – hidden columns, rows, and sheets; comments; and "track changes" (and similar in-line editing).

4.   When redaction is necessary, a redacted TIFF image version shall be produced. The Parties reserve the right to request access to the Native Format versions of such files.

B. Multimedia files

1.   Multimedia files (e.g., audio or video files) shall be provided in Native Format.

2.   In lieu of a TIFF image version of each multimedia file, a Bates-stamped single-page TIFF placeholder file shall be produced along with the Native Format version of each file. The TIFF Placeholder should read "Document Produced Natively."

3.   If a Producing Party wishes to redact multimedia files for reasons other than to protect the privacy of individuals not party to this Action, the Parties shall meet and confer as to the bases for such redaction, the format of redaction to be applied, and the format of production for the unprotected portion of the content. The details of any such redaction would then need to be included in a Privilege Log in accordance with the protocols set out in Section V of the Stipulation.

C. Other files

In limited circumstances, it may be necessary to obtain or view the Native Format versions of files, including color Documents/images and dynamic files, such as databases. The Parties reserve the right to request access to the Native Format versions of such files and/or an image format that supports viewing color Documents.

V.   **IMAGE LOAD/CROSS REFERENCE FILE REQUIREMENTS**

A. A single-page image load/cross-reference file shall be provided with each production.

B. The file should use the Concordance Image Viewer (".OPT") format, as in the sample below. Note, the volume label information ("MSC001" in the sample .opt file) is optional:

Sample Concordance Image Viewer .opt file:

MSC000001,MSC001,MSC\0000\00000001.TIF,Y,,,3

MSC000002,MSC001,MSC\0000\00000002.TIF,,,,

MSC000003,MSC001,MSC\0000\00000003.TIF,,,,

MSC000004,MSC001,MSC\0000\00000004.TIF,Y,,,2

MSC000005,MSC001,MSC\0000\00000005.TIF,,,,

VI.   EXTRACTED TEXT/OCR TEXT REQUIREMENTS

A. Extracted Text and/or OCR Text shall be provided for all Documents as separate, Document-level text files; Extracted Text and/or OCR Text shall not be embedded in the .DAT file. Document-level text shall be produced.TXT format for each Document, and filenames for .TXT files will be the Bates Number of the TIFF image of the first page of the corresponding Document.

B. OCR Text shall be provided for all redacted Documents in lieu of Extracted Text.

C. If a Document is provided in Native Format with a single-page placeholder TIFF image (e.g., spreadsheet files), the text file shall contain the full Extracted Text of the native file.

D. With respect to produced text files, the specifications herein address ANSI text-based productions. Unicode text-based productions shall be provided for productions containing non-English language Documents.

E. The Parties agree not to knowingly degrade the searchability of Documents as part of the Document production process. For example, Extracted Text should not be replaced by OCR Text except where appropriate (e.g., OCR for redacted Documents).

VII.   DATA LOAD FILE REQUIREMENTS

A. A data Load File shall be provided with each production.

B. The file shall be a Concordance-loadable data file ("*.DAT*" file) and contain Bates-

stamp and Metadata information as detailed below.

C. The delimiters and qualifiers to be used in the .DAT file are:

- *Record delimiter:* Windows newline/Hard return (ASCII 010 followed by ASCII 13)

- *Field delimiter:* Pilcrow ¶ (ASCII 020)

- *Multi-value delimiter:* Semicolon (do not follow with space) ; (ASCII 059)

- *Text qualifier:* Small thorn þ (ASCII 254)

- *Substitute Carriage Return or New Line in data:* Registered Trademark ® (ASCII 174)

D. With respect to .DAT files, the specifications herein address ANSI text-based productions. Unicode text-based productions shall be provided for productions containing non-English language Documents.

E. The .DAT file shall have a header line with the below listed field names and shall include the corresponding information per each applicable field:

Field	Comments	Document Types
Production:: Begin Bates	Beginning Bates Number.	All.
Production:: End Bates	Ending Bates Number.	All.
Production:: Begin Attach	Bates Number from the first page of each attachment.	Email only.
Production:: End Attach	Bates Number from the first page of each attachment.	Email only.
Production:: Image Count	Number of TIFF image pages in the produced Document.	All.
Number of Attachments	Number of attachments to an e-mail.	Email only.
File Extension	File extension of the original Document (e.g., .msg, .docx, .jpg).	All.
FileSize	File size of the original Document. Format: kilobytes.	All.
Unified Title	Subject of the document. If the document is an email, this field contains the email subject. If the document is not an email, this field contains the document's file name.	Non-Email files.

Field	Comments	Document Types
Custodian	Email: Mailbox in which the file was located. Other: Individual who originated the document. Format: LASTNAME, FIRSTNAME.	All.
All Custodian	For globally deduplicated productions only. Full name of all Custodians for whom the Document is being produced. Format: LASTNAME, FIRSTNAME; LASTNAME, FIRSTNAME.	All.
Author	Author of Document from available Metadata.	Loose files and attachments.
Email From	Electronic Communication author.	Electronic Communications only.
Email To	Electronic Communication addressee(s)/recipients.	Electronic Communications only.
Email CC	Email addressee(s), carbon copy.	Email only.
Email BCC	Email addressee(s), blind carbon copy.	Email only.
Email Subject	Email subject.	Email only.
Created Date/Time	File creation date. Format: MM/DD/YYYY HH:MM AM/PM.	All.
Last Modified Date/Time	File modification date. Format: MM/DD/YYYY HH:MM AM/PM.	All.
Sent Date/Time	Electronic Communication sent date. Format: MM/DD/YYYY HH:MM AM/PM.	Electronic Communications only.
Sort Date/Time	Primary date field MM/DD/YYYY HHMM AM/PM. This is the parent date copied down to attachments.	All.
MeetingStartDate	Start date of calendar appointment. Format: MM/DD/YYYY	Calendar items only.
MeetingStartTime	Start time of calendar appointment. Format: HH:MM:SS AM/PM	Calendar items only.
MeetingEndDate	End date of calendar appointment. Format: MM/DD/YYYY	Calendar items only.
MeetingEndTime	End date of calendar appointment. Format: HH:MM:SS AM/PM	Calendar items only.
TimeZone	The time zone in which files were standardized during processing (e.g., ET / GMT-5).	All.
MessageID	Globally unique identifier for a message assigned by the host email system. This value is extracted from parent message during processing and typically	Email only.

Field	Comments	Document Types
	includes messageid and a domain name (e.g., <0E6648D558F338179524D555@m1p.contoso.net).	
ConversationIndex	Email thread identification.	Email only.
Importance	Email flag indicating priority level set for message.	Email only.
Sensitivity	Sensitivity field from Email messages.	Email only.
Record Type	Descriptor for the type of document at the time of collection: "**Email**" for all e-mails; "**Attachment**" for files that were attachments to e-mails; "**Audio**" "**Video**" or "**Photograph**" for the applicable type of multimedia file; "**Electronic Communication**" for all files meeting the definition given in this Stipulation; "**E-Document**" for all other documents and files collected electronically; and "**Hard-Copy Document**" for Hard-Copy Documents that have been scanned and converted to an electronic image.	All.
FileName	Name of file as maintained in the ordinary course of business, at the time of collection.	Loose files and attachments.
Has Hidden Data	Denotes if file contains hidden content. Format: **Yes/No** value.	Loose files and attachments.
MD5 or SHA-1 hash	Unique file identifier.	All.
TextPath	The production deliverable path to the Extracted Text or OCR Text for the Document, including the file name.	All.
NativePath	The production deliverable path to the Native Format file for the Document, including the file name (if a Native Format file is provided).	Loose files and attachments.
ProdVol	Volume label information for the volume in which the document was produced.	All.
Confidential	Confidentiality level if assigned pursuant to the Protective Order. "**Whole**" for wholly confidential documents; "**Part**" for documents marked as partly confidential.	All.

F. In globally deduplicated, incremental productions, there will be instances when production of Documents from additional Custodians will include Documents previously

produced. A Custodian Append overlay Load File using the Load File format described above shall be provided with an updated AllCustodian field (BegBates and EndBates fields may be used as the unique identifiers). For globally and within custodian deduplicated productions, a similar append overlay Load File shall be provided with an updated AllFilePaths field.